

Considerations on Study Designs Using the Extreme Sibpairs Methods Under Multilocus Oligogenic Models

Chi Gu^{*,1} and D. C. Rao^{*,†}

^{*}Division of Biostatistics and [†]Departments of Psychiatry and Genetics, Washington University School of Medicine, St. Louis, Missouri 63110

Manuscript received April 2, 2001
Accepted for publication January 28, 2002

ABSTRACT

Several issues pertinent to study designs employing extreme sibpairs (ESP) methods to detect complex oligogenic quantitative trait loci (QTL) are investigated in the setting of genome-wide multipoint scans. We demonstrate that when stringent α -levels are imposed (*e.g.*, $\alpha = 0.00022$ as recommended by Landers and Kruglyak), the power to detect a susceptibility locus could drop from 83.6% under a one-locus model down to a hopeless 22.8% under a two-locus model of the same heritability ($h^2 = 0.5$) and gene frequency ($p = 0.1$). We introduce the notion of joint power that is the power to detect linkage to at least one location over a given panel of markers across a genomic region and describe the effect of several design factors on such joint power in a multipoint scan. Moreover, power of analysis conditional on the IBD sharings of ESPs at a known/detected locus is examined and shown to increase substantively (to 93.3% under the previous two-locus model) in detecting novel trait loci. We conclude that with such remedies, the ESP design continues to be a relatively powerful design for mapping oligogenic QTL. However, when the effect of individual contributing loci becomes less tractable, especially when their contributions are "asymmetric," deliberation on balancing two types of statistical errors and a careful examination of possible contributions from multiple genetic factors and/or interaction effects are a must in designing an efficient study.

As a powerful design in the case of quantitative traits influenced by a single gene, the extreme sibpairs (ESP) method can dramatically enhance the power by selecting sibpairs from extreme tails of the trait distribution (CAREY and WILLIAMSON 1991; FULKER *et al.* 1991; EAVES and MEYER 1994; RISCH and ZHANG 1995). However, it is well known that for complex traits, an increased number of quantitative trait loci (QTL) reduce power to detect linkage to any one of them, and it is not as clear how such selection techniques affect the power under multilocus oligogenic models. Possible epistasis among multiple loci may further reduce power to detect a particular QTL (CHEVERUD and ROTHMAN 1995). As a result, several problems arise immediately: (1) Given that the effect of an individual locus will be much less, should we still use a stringent α -level in genomic scans, (2) how do we minimize the effect of epistatic interaction on the power or utilize the information, and (3) is the effect of selective sampling still monotonic, and how does it affect the design of such a study?

This investigation is devoted to an examination of the power of ESP designs, when a quantitative trait is determined by two genetic loci. We show that although the effect of selective sampling is no longer monotonic,

the ESP method remains powerful when used with caution. We see that indeed the power under two-locus models is severely compromised if very stringent α -levels are imposed, making it necessary to strike a balance between false positives and false negatives (RAO 1998; RAO and GU 2001). To amplify the effect of a particular locus out of all possible epistatic interactions, we propose to conduct analysis at other sites conditional on the identity by descent (IBD) sharing at the known susceptibility loci and/or the loci where significant linkage was detected in earlier scans. Many reduced sample sizes are required when using such conditional analysis. Properties of similar conditional schemes for analysis of experimental crosses were thoroughly studied by Zeng and colleagues in their composite interval mapping (CIM) and multiple interval mapping (MIM) methods (ZENG 1994; KAO *et al.* 1999).

We begin with characterization of two-locus models using the notion of genotypic values *within* each trait locus. Trait distributions under several two-locus models are presented to demonstrate the increased challenge of devising selective sampling methods. Properties of two ESP statistics are then briefly reviewed. We then describe our method to calculate power of ESPs (either pointwise or jointly) by computing the probability of multipoint IBD distribution conditional on their trait outcomes and estimating the correlation structure of ESP test scores at any two marker locations.

Finally, we present power calculations over a variety of two-locus models and the effects of various kinds

¹Corresponding author: Division of Biostatistics, Washington University School of Medicine, Box 8067, 600 S. Euclid Ave., St. Louis, MO 63110. E-mail: gc@wubios.wustl.edu

TABLE 1

Genotypic means under an additive model

Genotypic values	B_1B_1	B_1B_2	B_2B_2
	$k + a_2$	$k + d_2$	$k + -a_2$
A_1A_1	$+a_1$	$k + a_1 + a_2$	$k + a_1 + d_1$
A_1A_2	$+d_1$	$k + d_1 + a_2$	$k + d_1 + d_2$
A_2A_2	$-a_1$	$k - a_1 + a_2$	$k - a_1 + d_2$

of ESP design. Important practical issues concerning oligogenic QTL mapping such as asymmetric contributions of loci, schemes of conditional analysis, utilization of different types of ESPs, threshold of trait selection, and overselection are carefully discussed.

METHODS

We assume that the quantitative phenotype X derives from an additive effect of the overall phenotypic mean (μ), several biallelic loci (g), and a residual term (e),

$$X = \mu + g + e,$$

where g is the genotypic mean conditioned on the genotypes of a person at the two susceptible loci. Two types of linkage tests using extreme sibpairs are discussed in this article. One was used by RISCHE and ZHANG (1995), which uses only one type of extreme sibpairs, discordant (ED) or concordant (EC), and is referred to as the ED test; the other one was introduced by GU *et al.* (1996), which combines both ED and EC [high-concordant (HC) or low-concordant (LC) or both] sibpairs and is referred to as the EDAC test. For a given sample size, the ED design is usually most powerful in the single-locus case, while the EDAC method is more cost effective. We see that under two-locus models the same principle applies. Namely, information from various types of sibpairs should be included to improve the power and the robustness of a study design.

Two-locus model: We consider, for the most part, two types of two-locus models in this article: an additive model with no interaction and an epistatic model with multiplicative interaction. Unless otherwise stated, the two contributing loci are assumed to have the same gene frequencies and contribute equally to the total heritability of the trait.

Let us denote the two alleles at the first major locus by a and A , with frequencies p_1 and $q_1 = 1 - p_1$, and alleles at the second locus by b and B with frequencies p_2 and $q_2 = 1 - p_2$. We assume that the alleles a and b contribute to increased risk of the disease/trait in consideration. The residuals are allowed to be correlated among relatives (correlation ρ). The genotypic means g , depending on the genotype of a person at the two loci, take values according to the underlying

TABLE 2

Genotypic means under a multiplicative model

Genotypic values	B_1B_1	B_1B_2	B_2B_2
	K_2b_2	K_2e_2	K_2
A_1A_1	K_1b_1	Kb_1e_2	Kb_1
A_1A_2	K_1e_1	Ke_1e_2	Ke_1
A_2A_2	K_1	Kb_2	Ke_2

$$K = K_1 \times K_2.$$

epistatic model as follows. Let $-a_1$, d_1 , and a be the “locus-specific” genotypic values for genotypes AA , aA , and aa , respectively, at the first locus and $k - a_2$, $k + d_2$, and $k + a_2$ be genotypic values for genotypes BB , bB , and bb , respectively, at the second locus, where d_1 and d_2 measure locus-specific dominant effects, and k indicates genetic displacement due to the second locus. The additive interaction between the two loci is defined in Table 1.

In a similar fashion, for the multiplicative model we use K_1 , K_1e_1 , and K_1b_1 as the “genotypic values” for genotypes AA , aA , and aa , respectively, at the first locus and K_2 , K_2e_2 , and K_2b_2 for genotypes BB , bB , and bb , respectively, at the second locus. And the multiplicative interaction is defined in Table 2, with $K = K_1K_2$.

The genotypic variance and the heritability under such models may be calculated as in KEMPTHORNE (1957), and their values depend on the gene frequencies, the genotypic displacement within each trait locus, and the magnitude of epistatic interaction of the two loci.

Selection of sibpairs on their trait values: The trait values are divided into a certain number of intervals with specified probabilities. Individuals with trait values that fall in either the top or the bottom intervals are said to have extreme-high or extreme-low trait values. An extreme sibpair has its members sampled from either one or both of the extreme tails.

We know that, under a single-gene model, the strength of the selective sampling method comes from the fact that by sampling sibpairs from the extreme tails of the trait distribution, the probability of their IBD sharing (2 or 0) is dramatically increased. It is only natural to expect that, under multilocus models, this enhancement will be weakened and the magnitude of the enhancement will be less clear and will depend much on the mode of interaction of the multiple loci. In Figure 1a, we plot the trait distribution of a two-locus model of additive gene interaction (no epistasis), assuming that both loci contribute the same to the trait (symmetric model) with $p_i = 0.20$, $k = 0$, $a_1 = 1$, and $d_i = 0$. We see that the effect of selection is nearly linear as in the one-locus case, except that its magnitude is smaller. In another case of a (symmetric) two-locus multiplicative model, as plotted in Figure 1b for $p_i = 0.2$, $b_i = 4$, $e_i =$

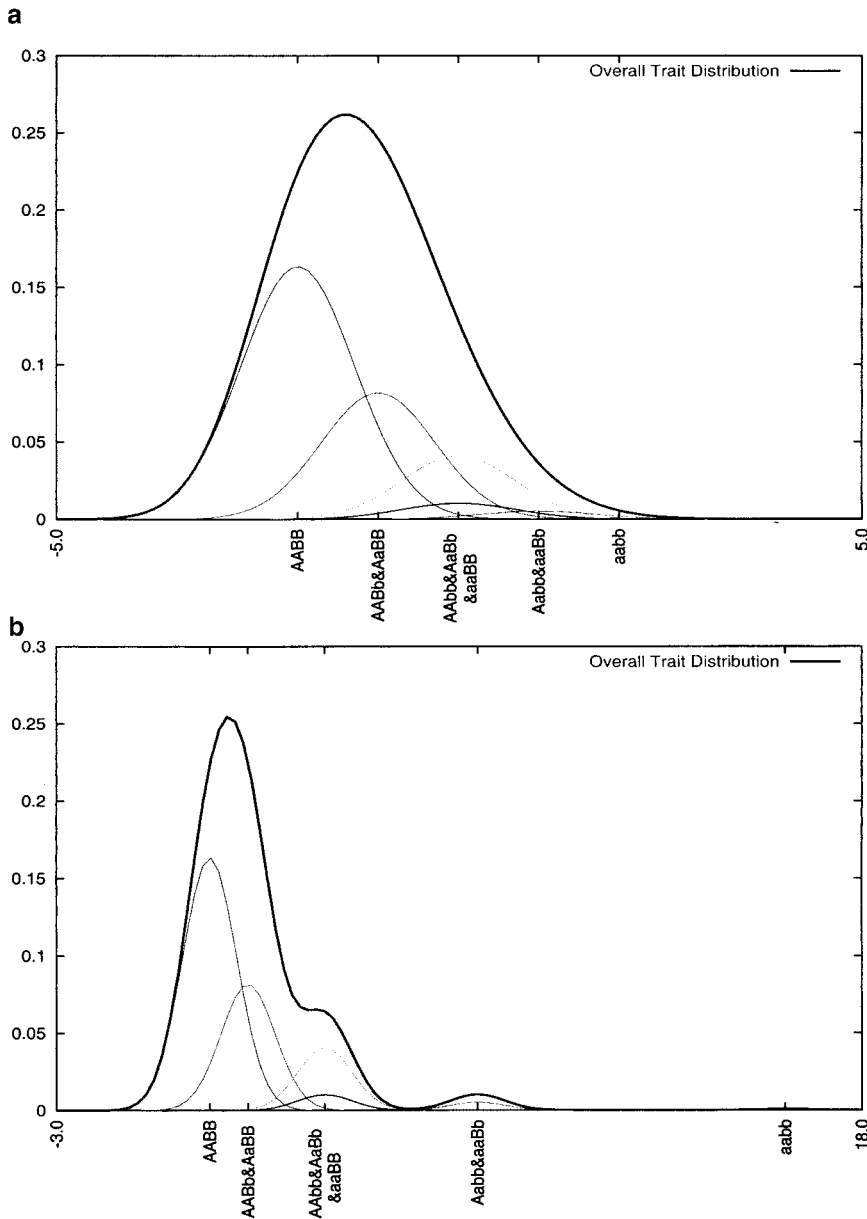


FIGURE 1.—Trait distribution of a symmetric additive model (a), both loci contribute the same to the trait with $p_i = 0.20$, $a_i = 1$, and $d_i = 0$, and a symmetric multiplicative model (b), $p_i = 0.2$, $b_i = 4$, $e_i = 2$, and $K = 1$.

2, and $K = 1$, the selection of high trait values becomes less effective because multiple distinctive genotypes linger near the upper extreme tail.

In Figure 2, we plot the trait distribution of an asymmetric two-locus model with additive gene action, for which $p_1 = 0.40$, $a_1 = 0.5$, $d_1 = 0$, $p_2 = 0.015$, $k = 2.0$, $a_2 = 2$, and $d_2 = 0$. The effect of selection on HC extreme sibpairs becomes more complicated since six genotypes for extreme sibpairs (including all possible genotypes at locus A) are now crowded in the upper extreme tail. Clearly the power to detect linkage to locus A will be much reduced compared to that to locus B, although both loci contribute the same heritability of 9.6% to the overall $h^2 = 19.3\%$.

Test statistics: We consider two types of ESP study designs. One is Risch-Zhang’s extreme sibpairs statistic (RISCH and ZHANG 1995), which uses a single type of

extreme sibpairs. Let k denote the number of sibs “affected” (*i.e.*, in the highest tail) in a sibpair (*e.g.*, $k = 1$ for ED pairs), π_{ki} be the proportion of genes shared IBD by the i th sibpair with k sibs affected, and n_k be the number of such sibpairs in the sample. If ED pairs are used, the Risch-Zhang statistic is given by

$$T_{ED} = \frac{1}{n_1} \sum_{i=1}^{n_1} \pi_{1i}$$

As an alternative, if there are sufficient numbers of both ED and HC sibpairs in the sample, the combined EDAC statistic (GU *et al.* 1996) is calculated as

$$T_{EDAC} = \frac{1}{n_2 + n_0} \left[\sum_{i=1}^{n_2} \pi_{2i} + \sum_{i=1}^{n_0} \pi_{0i} \right] - \frac{1}{n_1} \sum_{i=1}^{n_1} \pi_{1i}$$

Suppose that the statistics are calculated at each

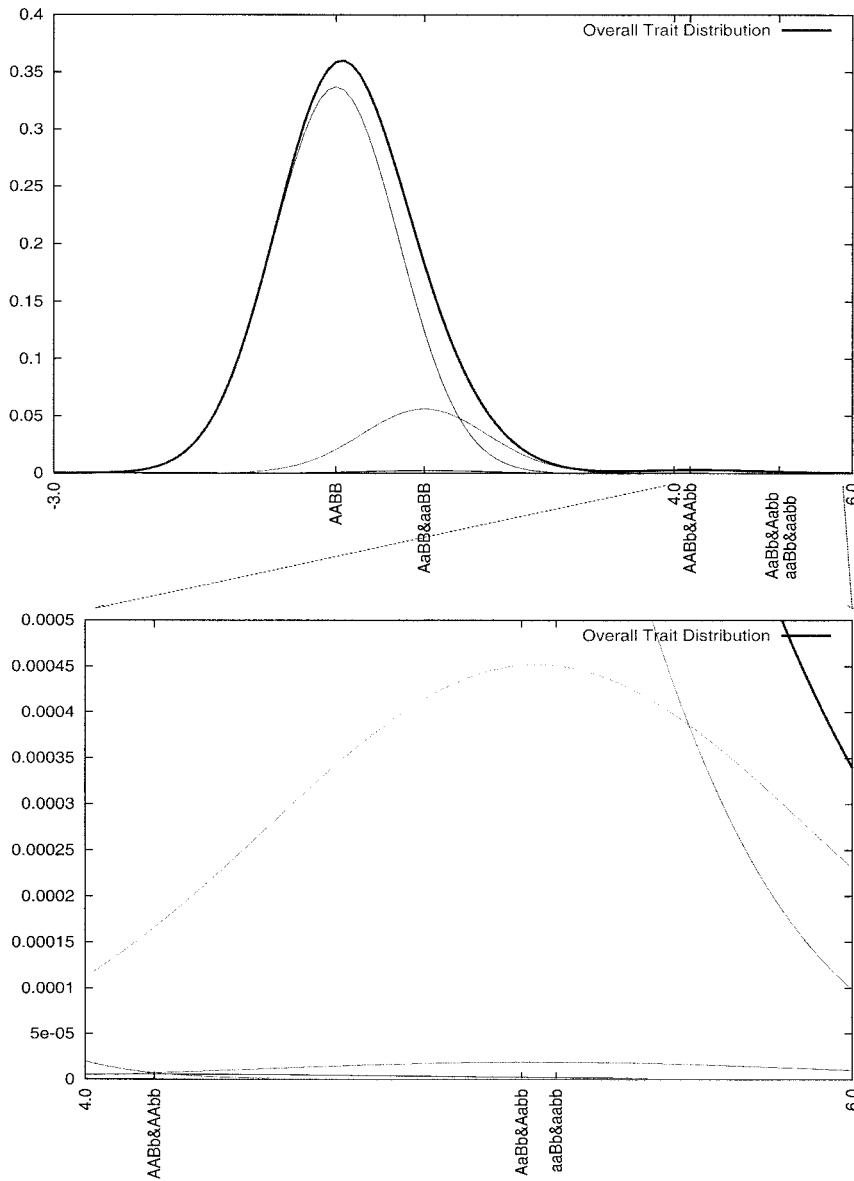


FIGURE 2.—Trait distribution of an asymmetric model (top) with additive gene action, $p_1 = 0.50$, $a_1 = 0.5$, $d_1 = 0$, $k = 2.0$, $p_2 = 0.016$, $a_2 = 2$, and $d_2 = 0$, and a magnification of its upper extreme tail (bottom).

marker in a panel of markers. We denote by X^k the statistic at the k th marker on the map. The power of ESP tests depends on, among other things, the expected IBD distribution conditional on phenotypes of a sibpair:

$$p_i = P(\pi = i|\mathbf{P}).$$

The formulas to calculate the power and/or sample size at a single marker locus are included below for your convenience.

For the ED test (see RISCH and ZHANG 1995), the power is

$$\Phi\left(\frac{|\tau - 0.5|\sqrt{2n_2} - 0.5Z_\alpha}{\sqrt{\tau(1 - 2\tau) + p_2}}\right), \tag{1}$$

where Φ is the cumulative distribution function of the standard normal, and $\tau = (2p_2 + p_1)/2$. The necessary

number of ED sibpairs to achieve a power of $1 - \beta$ is given by

$$n_1 = \frac{1}{2} \left(\frac{Z_{1-\beta}\sqrt{\tau(1 - 2\tau) + z_2} + Z_\alpha/2}{(\tau - 0.50)} \right)^2.$$

For the EDAC test (GU *et al.* 1996), the power is calculated by $\Phi(Z_{1-\beta})$, where

$$Z_{1-\beta} = \frac{E(T) - \sqrt{(n_1 + n_2 + n_0)/2n_1(n_2 + n_0)}(Z_\alpha/2)}{\sigma(T)}. \tag{2}$$

To calculate the necessary sample size for applying the EDAC test for a given power of $1 - \beta$ and significance level α , we first determine the ratios (r_0 , r_2) between the numbers of ED pairs and the expected numbers of LC and HC pairs in the sample. We then

TABLE 3

Power of 100 ED pairs under two-locus models, at the marker on top of the trait locus A, compared with that under one-locus models with the same heritabilities

h^2	$\alpha =$			
	0.01	$7.4e - 4$	$1.0e - 4$	$2.2e - 5$
0.10	0.039/0.113 ^a	0.005/0.019	0.001/0.005	0.000/0.001
0.20	0.132/0.504	0.024/0.194	0.006/0.078	0.002/0.036
0.30	0.348/0.918	0.104/0.685	0.035/0.462	0.015/0.313
0.40	0.648/0.999	0.308/0.978	0.145/0.919	0.075/0.836
0.50	0.866/1.000	0.583/1.000	0.360/0.999	0.228/0.994

The two-locus models assume additive gene action, with two *unlinked* loci contributing equally to the trait, and $p = 0.1$, $\rho = 0$.

^a Two-locus model/one-locus model.

substitute n_0 and n_2 by $n_0 = r_0 \cdot n_1$ and $n_2 = r_2 \cdot n_1$ and calculate the required number of ED sibpairs for the given power as shown in Gu *et al.* (1996), and sample sizes of LC (n_0) and HC (n_2) pairs are then computed using their respective ratios to n_1 .

Calculating the power of multipoint ESP scans: Under the multilocus/multipoint setting, to correctly calculate the power of the ESP tests, one needs to get the correct estimate of the expected IBD proportions across a set of markers conditional on the quantitative trait values of the sibpairs. The expected IBD sharing can be expressed as a sum of the products of the IBD probabilities at the *trait loci* conditional on the phenotypic values and the IBD probabilities at the *marker loci* conditional on the IBD sharings at the trait loci,

$$P(\boldsymbol{\pi}_m | \mathbf{P}) = \sum_{\boldsymbol{\pi}_t} P(\boldsymbol{\pi}_m | \boldsymbol{\pi}_t) \times P(\boldsymbol{\pi}_t | \mathbf{P}),$$

where $P(\boldsymbol{\pi}_t | \mathbf{P})$ is calculated by assuming that the trait values of a sibpair follow a bivariate normal distribution. The conditional probability $P(\boldsymbol{\pi}_m | \boldsymbol{\pi}_t)$ may be calculated

TABLE 4

Comparison of power ($\alpha = 0.0001$) under two-locus additive models, with two *unlinked* loci contributing equally to the trait: $p = 0.1$, $\rho = 0$, $N_{ED} = 100$

h^2	$\alpha =$			
	Locus A ^a	A and B ^b	Whole panel ^c	Select on B ^d (IBD = 2)
0.10	0.001	0.002	0.004	0.001
0.20	0.006	0.012	0.018	0.009
0.30	0.035	0.069	0.086	0.090
0.40	0.145	0.271	0.299	0.527
0.50	0.360	0.603	0.626	0.973

^a Pointwise power at locus A.

^b Joint power by scanning loci A and B.

^c Joint power by scanning the whole panel of 10 markers.

^d Pointwise power at locus A conditional on IBD = 2 at locus B.

TABLE 5

Comparison of power ($\alpha = 0.000022$) under two-locus additive models, with two *linked* loci contributing equally to the trait: $p = 0.1$, $\rho = 0$, $N_{ED} = 100$

h^2	$\alpha =$			
	Locus A ^a	A and B ^b	Whole panel ^c	Select on B ^d (IBD = 2)
0.10	0.001	0.003	0.006	0.862
0.20	0.015	0.030	0.045	0.948
0.30	0.108	0.193	0.237	0.984
0.40	0.399	0.604	0.651	0.996
0.50	0.752	0.920	0.942	0.999

^a Pointwise power at locus A.

^b Joint power by scanning loci A and B.

^c Joint power by scanning the whole panel of 10 markers.

^d Pointwise power at locus A conditional on IBD = 2 at locus B.

by the chromosome-based IBD distribution (CB-IBDD) algorithm discussed in Gu *et al.* (1995). Under the assumption of no interference, for example, unconditional probability of IBD sharing is given by

$$p(\boldsymbol{\pi}) = \frac{1}{2^4} \sum_s \prod_{k=1}^{m-1} [\theta_k^s (1 - \theta_k)^{2-s_k} - (1 - \theta_k)^s \theta_k^{2-s_k}] \times [\theta_k^k (1 - \theta_k)^{2-k} - (1 - \theta_k)^k \theta_k^{2-k}].$$

Consider the union map of both marker loci and trait loci. We first calculate probabilities of all IBD configurations with IBD sharing at the trait loci fixed by $\boldsymbol{\pi}_t$. And then we derive the conditional probability $P(\boldsymbol{\pi}_m | \boldsymbol{\pi}_t)$ as a proportion to the total sum of all such probabilities.

The power to detect linkage under multilocus models can be measured in two ways. One is the pointwise power, *i.e.*, the power to detect linkage at a particular marker location, which is calculated by Equations 1 and 2. The other is the total power of the scan over a panel

TABLE 6

Power of combined extremely discordant (EDAC) sibpairs, compared with that of ED pairs only, under two-locus models

h^2	$\alpha =$			
	0.01	$7.4E - 4$	$1.0E - 4$	$2.2E - 5$
0.10	0.105/0.056 ^a	0.018/0.007	0.004/0.001	0.001/0.000
0.20	0.483/0.226	0.183/0.053	0.074/0.015	0.034/0.006
0.30	0.896/0.577	0.651/0.247	0.435/0.108	0.295/0.053
0.40	0.995/0.888	0.957/0.624	0.876/0.399	0.779/0.260
0.50	1.000/0.986	0.999/0.898	0.992/0.752	0.977/0.611

Sample sizes of extremely concordant pairs are estimated as expected in the sampling pool for a given number of ED pairs. The two-locus models assume additive gene action and both loci contributing equally to the trait, with $p = 0.1$, $\rho = 0$, and $N_{ED} = 100$.

^a EDAC/ED only.

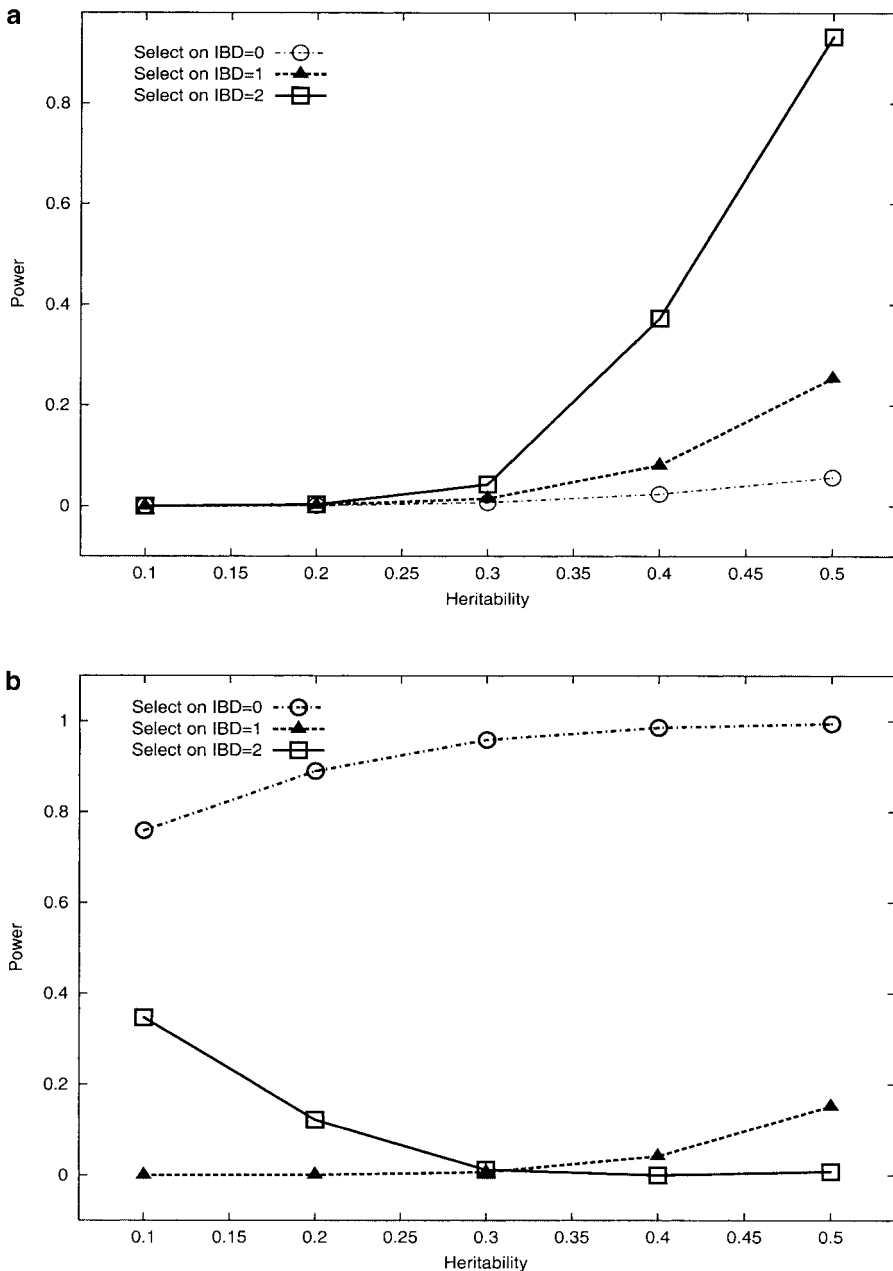


FIGURE 3.—Effect on the power of conditional analysis at locus A by conditioning on different IBD values at locus B. The two loci contribute equally to the trait under additive model with $p_1 = 0.10$ and $\rho = 0$. Results are shown when the loci are (a) unlinked and (b) linked.

of markers (called joint power in the sequel), *i.e.*, the power to detect linkage at least at one of the marker locations. We assume that the joint distribution of the ESP test scores across all the marker locations follows a multivariate normal distribution and estimate the covariance between the test scores at two marker locations as

$$\begin{aligned} \text{COV}(X^i, X^j) &= E(X^i X^j) - E(X^i)E(X^j) = 4(P_{22} - P_2 P_2') \\ &\quad + 2[(P_{21} - P_2 P_1') + (P_{12} - P_1' P_2)] \\ &\quad + (P_{11} - P_1 P_1'), \end{aligned}$$

where P_{kl} is the probability that a pair shares k genes IBD at locus i and l genes IBD at locus j , P_k is the probability that the pair shares k genes IBD at locus i , and P_l' is the probability that the pair shares l genes

IBD at locus j , and all are conditional on the sibpair's phenotypes. The formula involves only joint probabilities of IBD vectors and its calculation was described at the beginning of this section.

Conditional analysis: It is quite common in analyses of complex traits that one would detect, in a genome-wide scan, a number of *weak-to-moderate* signals among which many are possibly false positives (RAO and GU 2001).

One way to enhance the weak (but real) signals out of their noisy background is via signal amplification by conducting analysis at the interesting loci conditional on results at the other “candidate” or “detected” loci. For example, Cox *et al.* (1999) performed allele-sharing analysis of their diabetic families at markers across chro-

TABLE 7

Sample size of ED pairs using conditional analysis compared with that using unconditional analysis under two-locus multiplicative models

h^2	$\alpha =$			
	0.01	$7.4E - 4$	$1.0E - 4$	$2.2E - 5$
	$\rho = 0$			
0.10	58/3938 ^a	89/6019	113/7566	130/8716
0.20	49/883	75/1351	95/1698	110/1957
0.30	42/350	64/536	81/674	94/777
0.40	36/177	55/271	70/341	81/393
0.50	31/105	49/161	62/203	71/235
	$\rho = 0.4$			
0.10	45/541	70/827	89/1041	102/1199
0.20	35/153	54/235	68/296	79/341
0.30	29/76	44/117	56/148	65/170
0.40	25/48	38/75	49/94	56/109
0.50	22/36	35/56	44/71	51/82

Sample sizes are calculated to have a pointwise power of 90% at listed significance levels (α) at locus A. The two-locus models assume that two *linked* loci are contributing equally to the trait with allele frequency $p = 0.1$.

^a Conditional (on IBD = 0 at locus B)/unconditional.

mosome 15 weighted by the evidence for linkage at NIDDM1 on chromosome 2 and obtained a much clearer signal at the CYP19 locus (LOD score of 4.0 *vs.* 1.3; see Cox *et al.* 1999). We consider here the type of conditional analysis on the basis of the IBD status at

TABLE 8

Sample size of ED pairs using conditional analysis compared with that using unconditional analysis under two-locus additive models

h^2	$\alpha =$			
	0.01	$7.4E - 4$	$1.0E - 4$	$2.2E - 5$
	$\rho = 0$			
0.10	55/2366 ^a	85/3617	108/4547	124/5238
0.20	45/517	70/791	88/994	102/1146
0.30	38/202	59/310	74/390	86/450
0.40	33/104	51/160	64/201	75/232
0.50	29/65	46/100	58/127	67/147
	$\rho = 0.4$			
0.10	45/511	69/781	101/983	138/1133
0.20	33/131	52/202	76/254	104/293
0.30	28/65	43/100	63/127	87/147
0.40	24/43	38/67	56/85	77/99
0.50	23/34	35/54	52/68	72/79

Sample sizes are calculated to have a pointwise power of 90% at listed significance levels (α) at locus A. The two-locus models assume that two *linked* loci are contributing equally to the trait with allele frequency $p = 0.1$.

^a Conditional (on IBD = 0 at locus B)/unconditional.

TABLE 9

Effect of polychotomization on power under symmetric two-locus models is similar to that under single-locus models

T_L (%) ^a	T_H =			
	95%	90%	80%	70%
5	267	458	973	1801
10	303	518	1097	2024
20	375	641	1350	2480
30	462	790	1659	3040

Displayed are sample sizes of ED pairs required to achieve a pointwise power of 90% ($\alpha = 0.0001$) at locus A. The symmetric two-locus models assume additive effect at both loci and additive gene action, with both loci contributing equally to the trait. $p = 0.10$, $\rho = 0$, and $h^2 = 0.264$. More extreme thresholds result in more powerful tests.

^a T_L is lower threshold and T_H is higher threshold for defining extreme trait values.

“known” or detected loci. Namely, depending on the type of extreme sibpairs used in the analysis, we calculate test scores at other locations using only those extreme sibpairs that have specific IBD sharings at the detected/known genes. In other words, we zero-weight the extreme sibpairs with unfavorable IBD sharings at the detected/known genes in the calculation of ESP test scores at other locations. It is shown that conditional on the IBD status, the power can be increased substantially.

RESULTS

Power and/or required sample sizes are estimated under a variety of two-locus models. For fixed gene frequencies, fixed mode of epistatic gene action, and fixed ratio of genotypic values within each locus, we let the genotypic value vary to achieve a given overall heritability. Pointwise power, joint power over the two-trait gene locations, and joint power over a panel of 10 markers are all calculated. Also, cases of both *linked* and *unlinked* trait loci are considered. In the case of linked loci, the markers are assumed to be 10 cM apart and the trait loci are located at the fourth and the seventh markers. In the case of unlinked loci, the positions of the loci and the markers remain the same but the first and the last five markers are assumed to be from two unlinked groups (different chromosomes).

Balancing power and false-positive rate under multilocus models: For 100 ED sibpairs sampled from the top and bottom 10th percentiles of the trait distribution, we calculated the power to detect linkage at the marker on top of the trait locus A, the joint power to detect linkage at either A or B, and the joint power to detect linkage by scanning the panel of 10 markers. Displayed in Table 3 is the power to detect linkage under single-

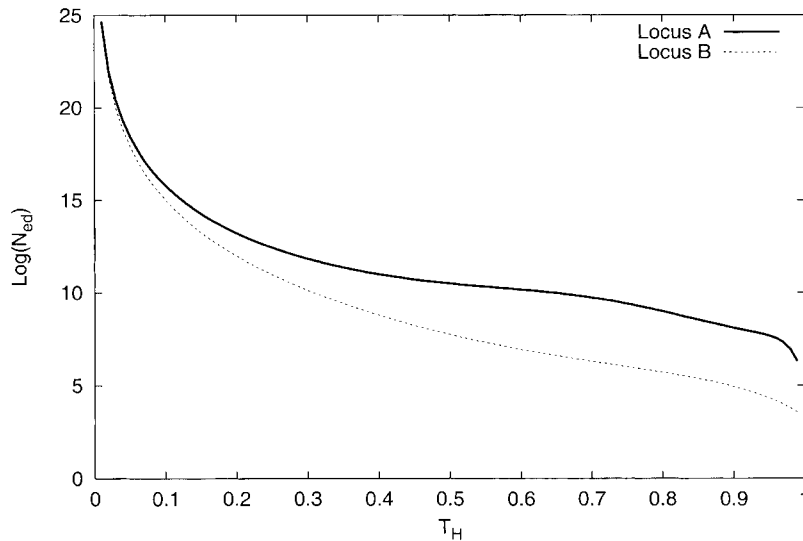


FIGURE 4.—Under a symmetric model ($p_i = 0.1$, $a_i = 1.0$, $d_i = 0$, and $k = 2$), required sample sizes of HC pairs for a power of 90% at $\alpha = 0.0001$ decrease at both loci A and B as the selection threshold of extremely high trait values increases.

locus models compared with that under two-locus models, with various preset false-positive rates (α -levels). It is worth noting that not only is the pointwise power severely reduced under two-locus models, but also more stringent significant levels reduced power drastically under two-locus models. Similar results hold when the two loci are linked.

Even when the overall heritability is fairly high ($h^2 = 0.50$), under two-locus models, the pointwise power is close to useless (always $< 23\%$) if the stringent $\alpha = 0.000022$ (LANDERS and KRUGLYAK 1995) is used to infer “significant linkage.” However, the joint power increases to 41.2% if two tests were performed on the markers right on top of the two trait loci. Using a less stringent $\alpha = 0.0001$ (corresponding to one false positive in ~ 18 scans with 400 markers), the power is $\sim 60\%$ if the two loci are unlinked and 92% if linked (Tables 4 and 5). As shown in columns 2–4 of Tables 4 and 5,

scanning multiple markers does improve power, and a couple of markers that are right on top of the two loci are as powerful as scanning a panel of 10 markers placed 10 cM apart. When the two trait loci are linked, scanning multiple markers can still achieve moderate power for somewhat lower heritabilities—if a less stringent threshold of significance is used. For example, under the previous model, one achieves a power of 70% at $\alpha = 0.01$ for $h^2 = 25\%$.

Enhancing power by EDAC and conditional analysis:

As we showed in previous work on ESP methods (GU *et al.* 1996; GU and RAO 1997a,b), combining different types of extreme sibpairs in the sample could significantly enhance the power to detect linkage. And for a variety of single-locus models, it is also a more cost-effective design compared with pursuing only a single type of extreme sibpairs. Our analysis verifies that similar conclusions hold for a variety of two-locus models.

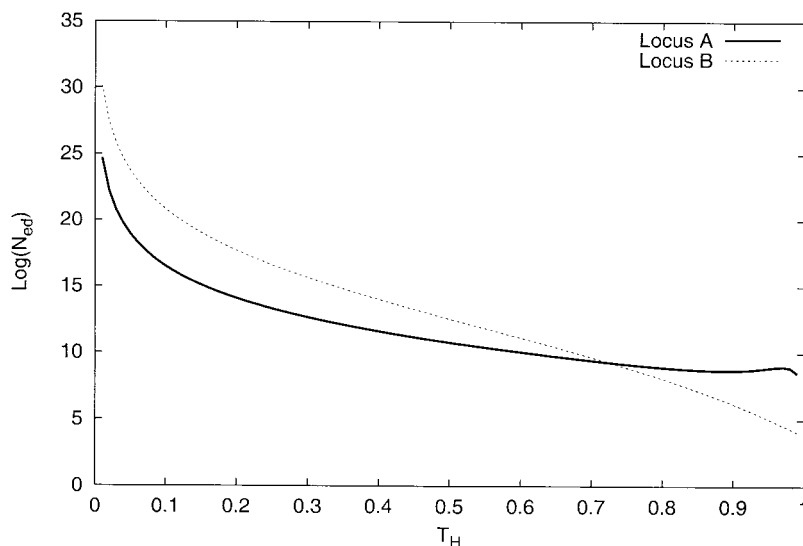


FIGURE 5.—Nonlinear effect of polychotomization on required HC sample sizes under an asymmetric additive model ($p_1 = 0.4$, $a_1 = 0.5$, $d_1 = 0$, $p_2 = 0.015$, $k = 2.0$, $a_2 = 2$, and $d_2 = 0$; see Figure 2 for trait distribution). The model has an additive effect at both loci, with a total heritability of 19.3%.

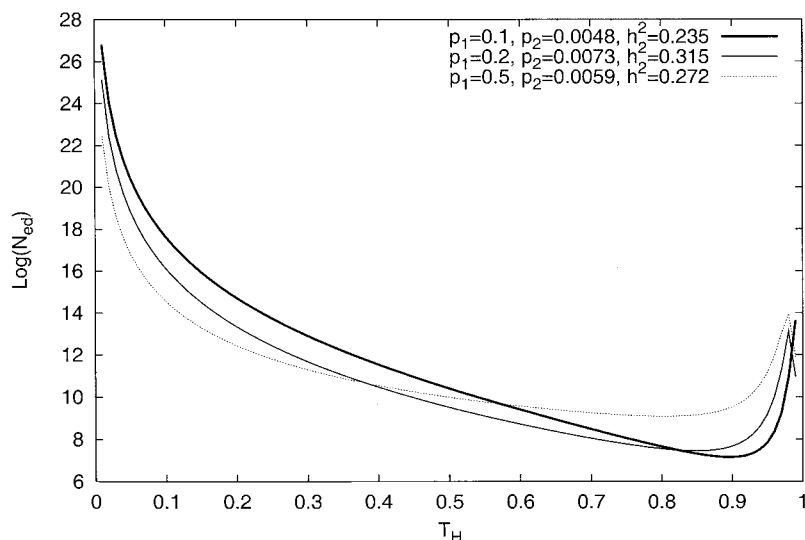


FIGURE 6.—Nonlinear effect of polychotomization on required HC sample sizes under asymmetric additive models ($a_1 = 0.5$, $d_1 = 1$, $k = 2.0$, $a_2 = 2$, and $d_2 = 1$ with various $p_1 = 0.1, 0.2$, and 0.5). The gene frequencies and the total heritabilities for each curve are displayed in the inset.

Displayed in Table 6 is pointwise power at the location of locus A calculated for 100 ED pairs under a symmetric additive model with $p = 0.1$, in contrast to the power of EDAC tests that also utilize HC and EC pairs available in the samples. The EDAC method consistently enhances power. For example, at the significance level of $\alpha = 0.00074$, the power of EDAC is 95.7% as compared to 62.4% for the ED-only design ($h^2 = 0.4$).

The enhanced power by conditional analysis is shown in column 5 of Tables 4 and 5. Note that since we used a ED-only ($n_{ED} = 100$) design the strategies used in the conditional analyses were different depending on whether the underlying disease loci were linked or not. Figure 3 shows the effect on power of different selection strategies. It is clear that for linked loci, selection on IBD = 0 and IBD = 2 resulted in very different power as the trait heritability increases. Assuming a symmetric multiplicative model with two linked loci, in Table 7 we display numbers of ED sibpairs required to achieve a

pointwise power of 90% at various significance levels to detect linkage to locus B using unconditional analysis, compared with that when the conditional analysis is performed by requiring extreme sibpairs sharing IBD of 0 at locus A. Sample sizes are reduced substantially using conditional analysis and the reduction is more substantial when heritability is small. Similar gain in power also holds under additive models as shown in Table 8.

Effect of polychotomization: The thresholds of trait values used to select extreme sibpairs are of importance in determining the actual power and cost-effectiveness of a QTL study (see GU and RAO 1997b for a discussion on determining optimal thresholds). Under multilocus models, the effect on power of polychotomization is no longer monotonic, nor is it comparable among trait loci, especially when the genetic contributions of the two trait loci are severely asymmetric. Nevertheless, over a variety of two-locus models our calculations show that ESP methods enhance power to detect linkage (as long as one maintains balance between the detection of either or both of the trait loci in the same scan and avoids overselection under “severely asymmetric” models, which may result in “unexpected” power decay).

In Table 9, we display the effect of polychotomization by calculating the required ED sample sizes for a power of 90% at $\alpha = 0.0001$ using different thresholds to select extreme sibpairs, under a symmetric model with additive gene action and additive effect at both loci ($p_i = 0.1$, $a_i = 1.0$, $d_i = 0$, and $k = 0$; $\rho = 0$, $h^2 = 0.264$). For a slightly different symmetric model ($p = 0.2$, $k = 2$), we show the required sample sizes of HC pairs to detect loci A and B, respectively, in Figure 4. Both examples demonstrate that, under the symmetric models, the power increases with more extreme thresholds for selection.

Under asymmetric models, however, the monotonicity between power and the selection threshold is likely

TABLE 10

Asymmetry of power to detect linkage at the two trait loci A and B, under asymmetric two-locus models

T_L (%) ^a	$T_H =$			
	95%	90%	80%	70%
5	618/3073 ^a	952/15649	1960/89590	3610/280992
10	697/3153	1068/15996	2185/91134	4008/285034
20	858/3289	1304/16580	2640/93723	4812/291806
30	1056/3418	1592/17137	3193/96190	5786/298255

Displayed are sample sizes of ED pairs required to achieve a pointwise power of 90% ($\alpha = 0.0001$) at A and B, respectively. The symmetric two-locus models assume dominant effect at each locus and additive gene action, with asymmetric gene frequencies and genotypic values: $p_1 = 0.1$, $a_1 = 0.5$, $d_1 = 1$, $p_2 = 0.005$, $k = 2.0$, $a_2 = 2$, and $d_2 = 1$. Each locus contributes the same to a total heritability of 0.235.

^a Required sample size to detect locus A/locus B.

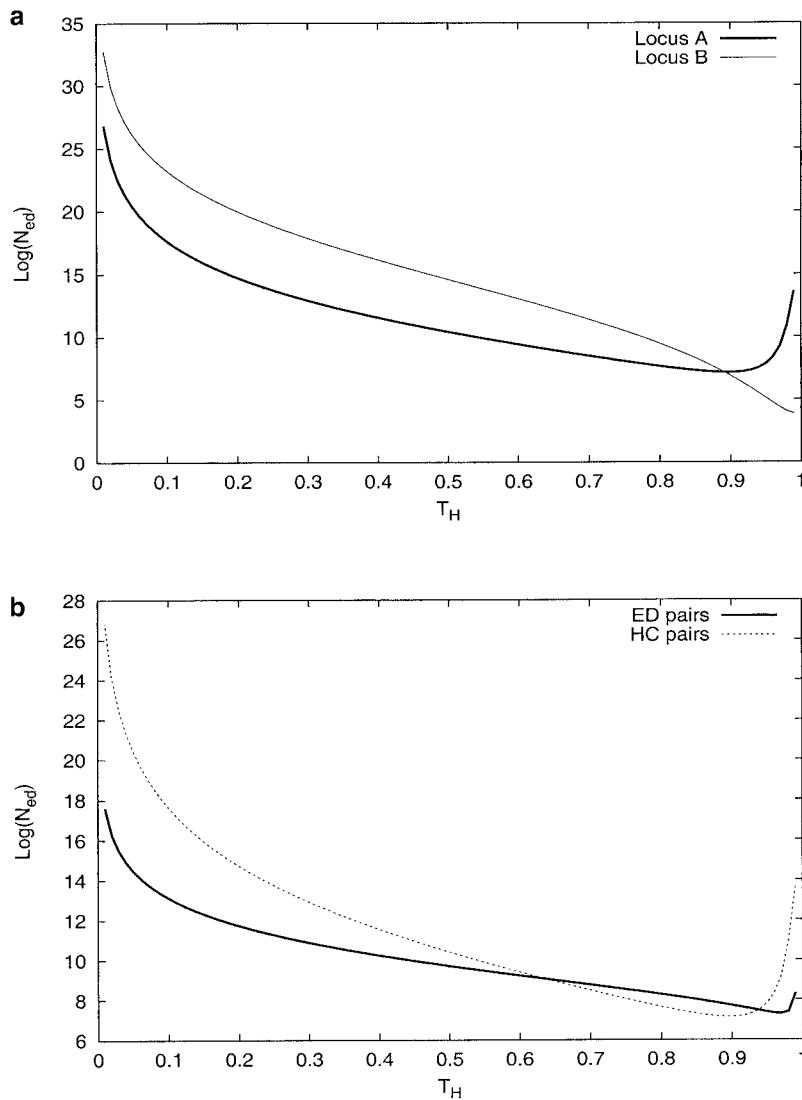


FIGURE 7.—Comparison of effect on power due to selection on trait values: (a) power at locus B *vs.* power at locus A; (b) ED design *vs.* HC design. The asymmetric model is the same as used in the previous example, with $p_1 = 0.10$, $p_2 = 0.005$, and the extremely low trait values are defined by a percentile of $100 \times (1 - T_H)/2$.

violated at one locus while maintained at the other, and the power to detect the two loci may not be comparable. For the model used in Figure 2 ($p_1 = 0.40$, $a_1 = 0.5$, $d_1 = 0$, $p_2 = 0.015$, $k = 2.0$, $a_2 = 2$, and $d_2 = 0$), we plot the required sample sizes of HC pairs (at the locations of A and B, respectively) for a power of 90% at $\alpha = 0.0001$ against the upper thresholds (percentiles) for extremely high trait values (Figure 5). We see that the power to detect A is slightly reduced when the threshold is >90 th percentile, but increased again when it is >98 th percentile.

For severely asymmetric models, overselection of extreme trait values could result in quite noticeable unexpected power reduction. Let us modify the previous asymmetric model so that it has a dominant effect at each locus ($d_1 = d_2 = 1$) and consider various gene frequencies at A: $p_1 = 0.1, 0.2$, and 0.5 . The required sample sizes of HC pairs to detect linkage at locus A are plotted in Figure 6. The gene frequencies at locus B are calculated so that it gives the same heritability as locus A. Under all the models, power increases with

selection threshold up to a certain level before declining unexpectedly. This shows that selecting for extreme trait values remains an effective strategy when moderately high thresholds are employed (see Table 10, where the power at loci A and B, respectively, is presented under several pairs of selection thresholds).

It seems that, under the models studied, higher gene frequencies at the problematic locus require less extreme selection to avoid power loss. The effect of selection on the power is also less significant in such cases (the curve in Figure 6 is flatter). It is worth noting that, in the above example, at locus B, the effect of selection on power is still proportional to selection (Figure 7a), and the power loss due to overselection occurs much later for ED pairs than for HC pairs (Figure 7b).

DISCUSSION

As the number of complex trait loci increases, it tends to weaken the linkage information carried at any given locus, even when the loci are closely linked. We have

shown that the extreme sibpair (ESP) methods remain powerful for detection of QTL of oligogenic traits in genome-wide scans. However, the power may be much reduced as compared to single-locus models and hopelessly so when stringent significance thresholds are used. Therefore, closer scrutiny is necessary when applying stringent significance thresholds in a genomic search. A balance must be maintained between the power to detect linkage and the accuracy of a linkage claim, *i.e.*, between false negatives and false positives (RAO 1998; RAO and GU 2001). For example, the α -level for infinitely many markers (continuous genome) and exceptionally low false-positive rate may not be the best choice. Moreover, in light of a recent article by RISCH and MERIKANGAS (1996), a candidate gene method may be the one to be used for fine-mapping of the trait genes. An extremely stringent significance level will often miss loci with small effects.

We provided a general algorithm for estimating the power of ESP methods under multilocus models and using multipoint analysis. It is clear that the power of an ESP design depends on the actual genetic model as well as the selection scheme for ESPs. The multilocus nature makes it harder to optimize study designs, since possible interaction among the trait loci and the possible asymmetry of all the genetic effects complicate the choice of a study design for detecting all the trait loci in a single genomic scan. As we have shown here, conditional analysis may improve the chances to delineate the epistatic interactions and detect linkage of smaller effects, by weighting test scores at certain locations according to the results at other locations already detected. We studied the type of conditional analysis that conditions on the IBD sharing of a sibpair at some detected/known genes. This may also be viewed as a type of subgroup analysis in which a subgroup of the sample is selected to enrich the homogeneity within the group. It will not only enhance the power of analysis but also assist in revealing interaction effects. Other weighting schemes may be applied in a similar fashion depending on the types of extreme sibpairs used and the actual results at the detected sites, and that certainly calls for further investigation. The ultimate solution to the problem of complex diseases, however, may depend on cleverly designed replication studies and the application of novel metaanalytical procedures tailored to pool genetic studies (GU *et al.* 2001).

This work was supported in part by National Institutes of Health grant GM28719 and National Institute of Mental Health grant MH31302.

LITERATURE CITED

- CAREY, G., and J. A. WILLIAMSON, 1991 Linkage analysis of quantitative traits: increased power by using selected samples. *Am. J. Hum. Genet.* **49**: 786–796.
- CHEVERUD, J. M., and E. J. ROTHMAN, 1995 Epistasis and its contribution to genetic variance components. *Genetics* **139**: 1455–1461.
- COX, N. J., M. FRIGGE, D. L. NICOLAE, P. CONCANNON, C. L. HANIS *et al.*, 1999 Loci on chromosomes 2 (middm1) and 15 interact to increase susceptibility to diabetes in Mexican Americans. *Nat. Genet.* **21**(2): 213–215.
- EAVES, L., and J. MEYER, 1994 Locating human quantitative trait loci: guidelines for the selection of sibling pairs for genotyping. *Behav. Genet.* **24**(5): 443–455.
- FULKER, D. W., L. R. CARDON, J. C. DEFRIES, W. J. KIMBERLING, B. F. PENNINGTON *et al.*, 1991 Multiple regression analysis of sib-pair data on reading to detect quantitative trait loci. *Read. Writ. Interdisc. J.* **3**: 299–313.
- GU, C., and D. C. RAO, 1997a A linkage strategy for detection of human quantitative-trait loci: I. Generalized relative risk ratios and power of sibpairs with extreme trait values. *Am. J. Hum. Genet.* **61**: 200–210.
- GU, C., and D. C. RAO, 1997b A linkage strategy for detection of human quantitative-trait loci: II. Optimization of study designs based on extreme sibpairs and generalized relative risk ratios. *Am. J. Hum. Genet.* **61**: 211–222.
- GU, C., B. K. SUAREZ, T. REICH and A. A. TODOROV, 1995 A chromosome-based method to infer IBD scores for missing and ambiguous markers. *Genet. Epidemiol.* **12**: 871–876.
- GU, C., A. A. TODOROV and D. C. RAO, 1996 Combining extremely concordant sibpairs with extremely discordant sibpairs provides a cost effective way to linkage analysis of QTL. *Genet. Epidemiol.* **13**: 513–533.
- GU, C., M. A. PROVINCE and D. C. RAO, 2001 Meta-analysis of genetic studies, pp. 255–272 in *Genetic Dissection of Complex Traits: Challenges for the Next Millennium*. Academic Press, San Diego.
- KAO, C.-H., Z.-B. ZENG and R. D. TEASDALE, 1999 Multiple interval mapping for quantitative trait loci. *Genetics* **152**: 1203–1216.
- KEMPTHORNE, O., 1957 *An Introduction to Genetic Statistics*. John Wiley & Sons, New York.
- LANDER, E., and L. KRUGLYAK, 1995 Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat. Genet.* **11**: 241–247.
- RAO, D. C., 1998 CAT scans, PET scans, and genomic scans. *Genet. Epidemiol.* **15**: 1–18.
- RAO, D. C., and C. GU, 2001 False positives and false negatives in genome scans, pp. 487–498 in *Genetic Dissection of Complex Traits: Challenges for the Next Millennium*. Academic Press, San Diego.
- RISCH, N., and K. MERIKANGAS, 1996 The future of genetic studies of complex human diseases. *Science* **273**: 1516–1517.
- RISCH, N., and H. ZHANG, 1995 Extreme discordant sib pairs for mapping quantitative trait loci in humans. *Science* **268**: 1584–1589.
- ZENG, Z.-B., 1994 Precision mapping of quantitative trait loci. *Genetics* **136**: 1457–1468.

Communicating editor: G. A. CHURCHILL

